

Generalization

Ki Hyun Kim

nlp.with.deep.learning@gmail.com

인공지능 모델이란?

- x 가 주어졌을 때, y 를 반환하는 함수

$$y = f(x)$$

- 파라미터(weight parameter, θ)란
 - f 가 동작하는 방식(x 가 들어왔을 때, 어떤 y 를 뱉어낼 것인가?)을 결정
- 학습이란
 - x 와 y 의 쌍으로 이루어진 데이터가 주어졌을 때, x 로부터 y 로 가는 관계를 배우는 것
 - x 와 y 를 통해 적절한 파라미터(θ)를 찾아내는 것
- 모델이란
 - 상황에 따라 알고리즘 자체를 이르거나, 파라미터를 이름

좋은 인공지능 모델이란?

- 일반화(Generalization)를 잘 하는 모델
- 보지 못한(unseen) 데이터에 대해서 좋은 예측(prediction)을 하는 모델
 - 우리는 모든 경우의 수에 대해서 데이터를 모을 수 없기 때문
 - 보지 못한 경우에 대해서, 모델은 좋은 판단을 내릴 수 있어야 함

기존 머신러닝의 한계

- 기존 머신러닝은 주로 선형 또는 낮은 차원의 데이터를 다루기 위해 설계되었음
- Kernel 등을 사용하여 비선형 데이터를 다룰 수 있지만, 한계가 명확함
 - 이미지, 텍스트, 음성과 같은 훨씬 더 높은 차원의 데이터들에 대해 낮은 성능을 보임